Woman C.P.A.

Volume 43 | Issue 2 Article 8

4-1981

Electronic Data Processing: Distributed, Interactive, & Integrated Systems

Elise G. Jancura

Advanced systems have been defined as those systems which possess one or more of the following characteristics: data communications, data integration, automatic transaction initiation, unconventional or temporary audit trail.[1] These characteristics are not unique to large-scale systems and in fact have been incorporated in many small processing.

Systems using data communications and data bases can operate in either a real-time or batch processing mode.[2] Most installations use a mixture of the two. Thus, an installation may use real-time processing for inquiries and batch processing for updating. Another approach would be to do all processing in real-time.

Data Communications and Remote Input/Output

Advanced systems frequently employ input/out devices which are at locations remote from the central processing unit. The remote devices are usually connected to the computer through telecommunication lines which are leased from a common carrier. The communication lines are available at various speed ratings and are usually either *dial-up lines* (the connection is completed by use of a normal dialing procedure) or *leased lines* (dedicated service, the user has exclusive use of the communication line). The remote locations of the input/output devices make it more difficult to control access to the system than in an installation where all input devices to the computer exist within the controlled environment of the physical installation itself. Thus, techniques have to be introduced to identify the users of the I/O devices and to control their access to the central processing system and its information files.

This need for control of access exists in any environment but is especially critical when input/output devices that have entry to the processing system are at locations physically remote from that of the central processing unit. It should be stressed that the use of remote input/output devices and telecommunication lines does not make a system a real-time system. If a group of transactions are collected at a remote location and transmitted at one time to a central location, this is a batch processing system. The characteristic that makes the use of these telecommunication lines part of a real-time system is recording, transmitting, and processing the transaction immediately as it occurs.

There are a great variety of devices available for use as remote input/output operations. Some of these devices are manually operated terminals in which the operator uses a keyboard or some other device, such as a light pen, to actually record the data for transmission, usually one record at a time. Other devices are essentially high-speed transmission devices which read prepared documents (cards, magnetic tape, or other machine-readable records) and usually transmit these documents a group or batch at a time. When such high-speed devices are in use, they are usually employed in some form of remote job entry and are usually involved in a batch processing application.

Some installations use telecommunication lines to communicate with remote locations in an off-line environment. In off-line activities, data is not transmitted directly to the computer but is transmitted to another device located at the installation, such as a magnetic tape handler or a card punch, and the machine records thus received are held for later processing. In other installations, the remote devices are linked through the communication lines directly with the central processing unit. In still other applications, the data communication lines can be used to link two computers which can communicate with each other. An increasingly popular configuration is one in which a network of small computers, usually used for local processing, can be linked to large central computers so that the sharing of information and processing can occur throughout the network thus created. In such configurations, conventional files are frequently replaced by data bases. This approach is called *distributed processing.*

The use of remote input/out devices in data communications is available in both large and small systems. It has the effect in many applications of moving the data capture function closer to the source of the data and the original user. This technology also makes it possible to move information from one processing system to another automatically.

Remote input/output devices and data communications do not change the requirements for physical protection of the equipment and data files. However, the existence of remote input/output devices by which access can be gained to the files of the system through long distance transmission lines does present a potential control problem. First, there are many more potential access points compared with the rela-

Electronic Data Processing

Distributed, Interactive, & Integrated Systems

Editor:
Elise G. Jancura, CPA, CISA, Ph.D
The Cleveland State University
Cleveland, Ohio 44115

tively few devices in a system without remote terminals, and second, the remote location of these input/output devices makes the physical control of the devices more difficult. Procedures for identification and authorization of users of the remote devices are critical.

Integrated Files and Data Base Systems

Historically, computerized accounting systems employed conventionally structured files— i.e., a separate file for each application or group of applications. This approach had the advantage of simplicity, for a given file contains all the data items needed for the programs that process it. However, it also has disadvantages:

Some of the information in one file is also contained in other files and thus results in data redundancy.

Several files may contain different versions of data common to both of them. This happens when not all files are updated by current transactions on a timely basis or when the various updating programs produce different results.

These disadvantages can be mitigated by integrating the data items into a few common files, perhaps even a single file, called a *data base.* These data remain on-line and are processed by all the programs in a given system. For example, there may be a data base for the sales, accounts receivable and cash receipts system, another for the purchases, accounts payable, cash disbursements system and yet another for the payroll-personnel system. Or, there may be one data base for everything: all systems access the same data base.

Data integration results in the combination of the data records for several different operations with similar information into single comprehensive sets of records. This process of creating single comprehensive records and thus a single comprehensive file minimizes the necessity for duplicate operations and duplicate records. These integrated sets of data records, called the data base, become the master file for a number of different applications in the system.

The processing for a data base is characterized by the fact that a single document describing a transaction is used to initiate the updating of all records or data elements associated with that transaction and affected by it. Although this results in an elimination of redundancy within the master files and more efficient handling of all facets of the transaction, it places a very heavy responsibility on the installation for maintenance of that single data base. Under this approach all of the pertinent master information and historical data is contained within one single master file or data base, and erroneous processing or inadvertent destruction of that single data base can have more serious implications for an organization than the destruction of an individual application master file that is only one of several master files for a firm. Data bases can be updated or integrated on either a real-time or a batch-mode basis.

Control Requirements for Integrated Files

Access controls for the data base should encompass both those activities that can affect a change in the files (updating) and those activities that simply involve reading the data files (inquiry). The data files represent a real asset whose value can be diluted as much by unauthorized access and reading as by incorrect updating. Access controls become even more critical when the data base is maintained on-line, making the information accessible to all devices within the computer system. This is to be contrasted with files that are maintained on an off-line basis, such as tape, and require an actual mounting of the data file before they can be interrogated through the computer. If the on-line data base exists in a system which has remote terminals, access controls must be concerned with all users having access to either the remote terminals or to the central processing facility.

In an integrated data base containing many different data elements from different applications, individuals may have authorization to access specific modules or elements within the data base. A carefully constructed system of authorization for access to each data element in the system must be established to prevent improper access or manipulation by persons without legitimate reason for accessing the information. Furthermore, responsibility for each data element in the data base must be established and assigned. A good security system must not only control which data elements a user is entitled to access but also the operations that may be performed. Control should be exercised over which programs a user is allowed to execute. Whether an existing record in the file may be updated or simply read, whether the user is authorized to add a new record, or whether the user is authorized to eliminate or delete a record from the data file must all be made a part of the system specifications and controls.

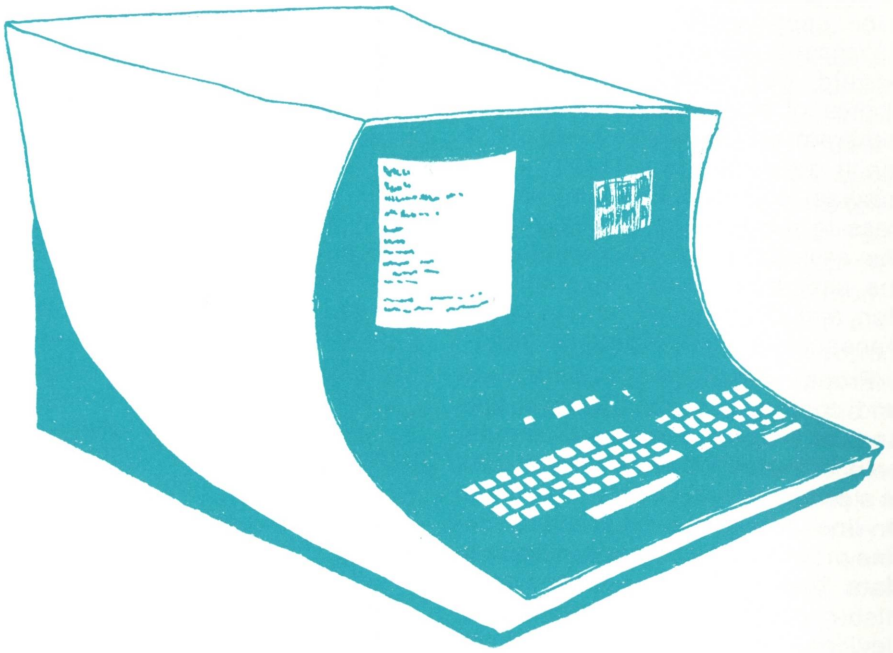
Another function that must be carefully controlled within the installation is the identification of those individuals who have the right to change the access authorizations within the system or execute those programs that interpret or otherwise update the user identification routines. Once a system of identification has been worked out so that the processing routines and the data files the individual user may access have been identified, it is important to protect the integrity of these identification codes or passwords. Individual users must be impressed with the importance of maintaining the confidentiality of their password, and steps must be taken to prevent unauthorized access to the identification scheme. One of the easiest ways to permit access to the passwork or identification code is to allow those passwords to be printed on a terminal during the normal processing performed on that particular terminal.

Passwords or identification codes should not be made available to the computer operators or systems programmers on a normal basis. Consideration should be given to periodic rearrangement or changing of the passwords in order to prevent publication of the passwords. Installation management must control access to the identification codes and processing programs to prevent systems programmers from either gaining access to that information or modifying it for their own use. The installation management should exercise continual care in reviewing the operations of the computer to insure that knowledgeable operators have not themselves attempted to access the data.

In addition to identifying legitimate users and preventing unauthorized access to the data files, a file security system should also provide regular notification to the operations management or the security officer of the installation when unauthorized attempts to access the data base are made. Attempts to gain unauthorized access can be handled immediately by techniques that lock the terminal keyboard or simply fail to respond to these initial inquiries. In the case of critical information, such an attempt could cause a message to be relayed to an appropriate security officer who could investigate the terminal in question.

In addition to preventing access, it is desirable to maintain a log of unauthorized attempts to access the data files in order to determine whether there is some pattern involving a particular user, a particular terminal location, or a particular element of the data base. If a pattern can be discerned, the opportunity to identify the illegal user or strengthen the security system is enhanced. In the case of a particularly sensitive element, it may be desirable to maintain a log of all accesses to the records within that data element. This log can then be periodically reviewed to assess the way that segment of the data base is being used as well as to review the efficiency of the control procedures maintained for that file.

As with any other data files, provision must be made to provide adequate backup for the data base system. Since data base systems are usually recorded on direct-access devices and processed by destructive updating, which does not provide a grandfather-father-son backup procedure, it is necessary to make some provision to prouce backup copies of the file periodically. This can be done either by periodic dumping or by an adaptation of a logging procedure. As each update takes place within the data base, a log can be constructed before processing the transaction to create a copy of the master record. After processing, a copy of the updated master record would be made. If any disruption of processing takes place or if the data base were physically damaged, it would be possible to use the latest file dump or the reconstruction log (whichever technique is used) to reconstruct the physical data base.



In a data base environment, the organization and maintenance of the master data is separated from the application programs. A separate group of software programs, called the *data base management system,* organizes, records, and retrieves data elements from the data base making specific elements available to the application programs when updating is necessary. This is in constrast to the approach used when individual application files are created and maintained by the application programs themselves. This separation of function causes some change in the approach to maintenance and protection of the data files. Many installations use a data base administrator whose responsibilities are to develop the organization of the data base, including identification of the elements and the logical relationships of the individual data elements to each other and to the various applications systems. The data base administrator is also responsible for the documentation of the data base and the implementation of the control and security measures developed for that data base system.

In organizations large enough to support staff specialization, the data base administrator represents a separate individual or staff. It should be recognized, however, that many small installations employing minicomputers also use the data base approach. While the separation of responsibilities makes it desirable that a different individual act as data base administrator from the person operating the system and developing application programs, that separation may not be possible. If this is the case, then alternative controls must be instituted to compensate for that lack of segregation of responsibilities within the staff.

Impact on the General Controls[3]

The use of data communications, on-line integrated files, and the availability of interactive processing generally results in a great deal of automatic processing of a transaction after its introduction into the system. This will increase the significance of the programmed systems controls and the controls dealing with the transaction-initiation function. Separation of the responsibilities for initiating transactions and processing them remains imperative as the data-capture function is moved closer to the user group through the use of remote terminals.

The use of a transaction log, controls totals, or internal program

checks that verify the accuracy of the transaction itself are all more meaningful when provision is made for independent input records in the user departments. This does not necessarily mean that the system should be forced to reproduce copies of source documents or to generate unnecessary printed matter. It does require that the system provide for carefully controlled access to the input terminals and for the assignment of responsibility to the user department for authorization and subsequent verification of transactions used to process data.

Proper identification of the user, and control of access to data files and the program libraries is another area of increased significance. This is a particularly sensitive issue in an on-line environment, featuring on-line program libraries and integrated data files in a processing system also containing multiple remote I/O devices. Because so much of the processing is automated, protection of the integrity of the program libraries—both application programs and systems support programs, such as the operating system—is critical. Logs should be maintained of all accesses to those program libraries and of any changes made to the libraries. Periodic tests should be made of the program library to insure that unauthorized changes have not in fact occurred. Authorized changes to the program library should be properly documented indicating the nature of the change and the authorization for the change.

The installation should be prepared to react to any unexpected emergency occurring within the processing environment. Correction routines should be in operation to handle normal errors within transmission or processing procedures. In addition, preplanned procedures and programs should be available to handle major failures and the resulting restart activities. Restarting a distributed system with multiple input/output operations or a real-time updating operation requires a sequence of operations that frequently is quite demanding. The potential for losing transactions or for repeating others is great, and the potential for operator error is increased under the duress of an emergency situation if the operator has not be thoroughly trained in a predetermined and tested plan of action for emergency situations. Similarly, good processing control requires that adequate attention be given to the need for file reconstruction and for the preservation of sufficient file and transaction information to allow for reconstruction of the data base in a acceptable time frame.

Once detected, error conditions within the system must be corrected. This may require reintroduction of corrected data. Error-correction procedures are complicated when the transactions are initiated at a remote location or when the updating is taking place in a real-time environment.

When data is introduced from remote locations, the system should be designed to notify the terminal user of the disposition of each transaction. Then should a system failure occur, the terminal user will be aware of those transactions that have been processed and of those transactions that have not been processed and require additional operator action such as reentry into the system.

Impact on Application Controls

The use of programmed self-checking digits, existence checks, combination checks, completeness checks, and reasonableness checks provides a method by which individual transactions can be examined for accuracy by the computer before the transaction is used in processing. If the processing is being done in a batch mode, these programmed checks can be implemented in an edit program which checks all of the transactions at one time. In a real-time environment these checking procedures must be included in the data entry program itself since the nature of real-time processing is to process a transaction as soon as it occurs. Again, it should be recognized that the nature of an application control is not changed by the existence of remote terminals or integrated on-line files. The decision to apply this technology in either a batch-mode or real-time environment, however, will affect the timing of some of these application controls.

Accumulative control totals, record counts, or logs of transactions continue to be effective control tools. In the case of batch processing environments, these accumulated totals can be used to balance data prior to the next processing step. In the case of real-time processing, these totals are not accumulated until after the processing has been completed. But they continue to be useful, for they provide a means to verify that all transactions have in fact been recorded and subsequently processed. In the real-time mode, as transactions are entered over a period of time, they can be posted to accumulative totals. At the end of the specified period of time, the accumulated control totals can be compared with similar controls being maintained at the initiation point. The balancing operation can be facilitated by procedures that identify the terminal and/or the user and can accumulate controls by these subdivisions. The presence of the balancing or control totals can be used both to verify application controls and to check on the transmission accuracy of the communications network. Remember that in a real-time environment these control totals or batch totals are used after the fact. They should always be used as a supplement to the data verification procedures implemented before the transaction is used in processing.

These control totals can usually be accumulated as a by-product of the transaction log. While the transaction log can be used for balancing and improving the total processing performed over a period of time, it usually provides no protection against the omission of original transactions; the log simply contains copies of original transactions actually entered through the input terminals. Further, if an erroneous transaction goes by the input controls and is allowed to enter the system, it would be reflected as it was entered on the transaction log. If independent controls are maintained on the actual transactions and the control totals (calculated either through an accumulation routine or totals from the transaction log) are compared with these independently developed control totals, a better control exists for determining that all transactions have in fact been entered into the system.

In those instances where the system itself generates a transaction (as in the automatic reorder function

in an inventory control system), it is important that the system document the existence of that machine-generated transaction by producing some hard-copy memorandum that can be verified by an independent check of the activity.

All transactions must be properly authorized by the user department. Listing all transactions processed during a preceding period and returning the list to the supervisor of the original department facilitates verification of the authorization of all processed transactions. Similarly, restrictions of terminals to authorized users can also constitute approval of the documents generated by that terminal.

Controls must also be instituted to prevent loss of transactions. The control totals mentioned previously represent one approach. Still another that can be employed is to provide serial numbers for each transaction. In some cases this can be done by the terminal operator. Thus, as each operator sends a message, a serial number is attached to the message. As the computer receives the message, it can check that the serial number from each message received is one higher than the serial number from the previous message received from that particular operator or that particular terminal. A variation on this approach is to allow the computer itself to generate a serial number which is attached to each transaction and recorded on the transaction log.

The use of serial numbers allows the system to keep track of all transactions entered into the system. If a hardware or software breakdown occurs, or if the application program malfunctions, the system has a means of identifying the fact that a transaction has not been properly processed. System-generated control or serial numbers can be used in the same way that serial numbers or preprinted numbers are used on normal source documents. For those computer systems that contain electronic clocks, the serial number can be expanded to include a time designation as well.

Summary

The techniques discussed above make systems potentially faster, more efficient, and more versatile. It is possible to move data collection closer to the source of transactions thereby eliminating a lot of previously necessary data handling activities. In some cases data acquisition can be handled automatically without any human intervention. The use of integrated files or data base systems can eliminate the duplication of information that previously existed in separate application files and can facilitate more efficient access and updating of those integrated files. These advantages make it possible to use automated systems in all aspects of an organization's operations and to make information available on a sufficiently timely basis to be an important resource in the decision-making process.

While these technological advances offer promise for increased efficiency and effectiveness, they also represent the need for new approaches to control and security. Communication techniques that allow for remote collection of data at the user site also produce the opportunity for unauthorized access to the processing system if adequate controls are not implemented to protect against that danger. Elimination of intermediate processing and automatic generation of transactions may eliminate previously available documents and thus change the traditional audit trail. In order to enjoy maximum benefit from the efficiencies of the techniques available in advanced systems, it is necessary to develop control procedures which will protect the integrity of these systems and the data processed.

NOTES

[1]*Management, Control, and Audit of Advanced EDP Systems,* (New York: American Institute of Certified Public Accountants, 1977), p. 5.

[2]For a discussion of real-time systems, see "Implications of Real-Time Systems for Accounting Records," by Elise G. Jancura in *The Woman CPA,* January 1975 and April 1975.

[3]See the Audit and Accounting Guide, *The Auditor's Study and Evaluation of Internal Control in EDP Systems,* (AICPA, 1977) for a detailed discussion of General Controls and Application Controls.